Neuroscience of Consciousness, 2016, 1–10

doi: 10.1093/nc/niw016
Research article

Unconscious associative learning with conscious cues

Andrea Alamia[1,*], Jean-Jacques Orban de Xivry[1,2], Estibaliz San Anton[3], Etienne Olivier[1], Axel Cleeremans[3] and Alexandre Zenon[1]

[1]Institute of Neuroscience, Université catholique de Louvain, Brussels, Belgium; [2]Department of Kinesiology, Movement Control and Neuroplasticity Research Group, KU Leuven, Leuven, Belgium; [3]Consciousness, Cognition, and Computation Group, Centre de Recherche Cognition et Neurosciences, Université Libre de Bruxelles, Brussels, Belgium

*Correspondence address: Institute of Neuroscience, Université catholique de Louvain, Brussels, Belgium. Tel: 02 764 54 32; E-mail: andrea.alamia@uclouvain.be

Abstract

Despite extensive research, the very existence of unconscious learning in humans remains much debated. Skepticism arises chiefly from the difficulty in assessing the level of awareness of the complex associations learned in classical implicit learning paradigms. Here, we show that simple associations between colors and motion directions can be learned unconsciously. In each trial, participants had to report the motion direction of a patch of colored dots but unbeknownst to the participants, two out of the three possible colors were always associated with a given direction/response, while one was uninformative. We confirm the lack of awareness by using several tasks, fulfilling the most stringent criteria. In addition, we show the crucial role of trial-by-trial feedback, and that both the stimulus–response (motor) and stimulus–stimulus (perceptual) associations were learned. In conclusion, we demonstrate that simple associations between supraliminal stimulus features can be learned unconsciously, providing a novel framework to study unconscious learning.

Key words: Unconscious learning; implicit learning; consciousness; motion discrimination; reinforcement learning

Introduction

Unconscious learning can be defined as "learning without awareness, regardless of what sort of learning is being acquired" (Shanks and Stjohn, 1994). One can frame the current literature on unconscious learning along two dimensions: the first one determines whether the stimuli used during learning are supraliminal or subliminal, whereas the second dimension characterizes the complexity of the rules or associations to be learnt (Fig. 1A). More specifically, the term "complex rule" is used here to refer to task structures that are composed of a large number of contingencies, such as in sequence learning tasks (in which numerous transition between successive key presses have to be learned), or artificial grammar tasks (in which a set of probabilistic rules drive the generation of grammatical strings). In opposition, simple rules can be defined as task structures composed of a small number of contingencies. Notably, there is not a precise separation between these two classes of rules, which rather define two extremes of a continuum. The use of supraliminal stimuli to induce learning of abstract, complex rules (top left corner in Fig. 1A) is the hallmark of implicit learning (Reber, 1967). Throughout the years, many experimental paradigms have provided reliable and replicable evidence of implicit learning, from artificial grammar learning (Reber, 1967) to sequence learning (Nissen and Bullemer, 1987), from control of complex systems (Berry and Broadbent, 1984) to statistical learning (Saffran *et al.*, 1999). In the example of the artificial grammar case, the complexity of the rules ensues from the presence of a large set of probabilistic associations that generate the strings. Moving to the opposite side of the theoretical space described in Fig. 1A, subliminal stimuli used to learn simple associations set the framework of what has been defined as subliminal learning (Clark and Squire, 1998; Clark *et al.*, 2002;

Received: 12 January 2016; **Revised:** 14 June 2016. **Accepted:** 10 August 2016

Olsson and Phelps, 2004). In a typical subliminal learning paradigm, a cue, which is nonconsciously perceived, predicts an outcome or prompts a response. In this case, the complexity is low since there is only one association driving the behavioral effect, although the stimuli are perceived subliminally. Looking at the top right corner of Fig. 1A, we find the learning of subliminal stimuli associated with complex rules. This has been poorly investigated because of both methodological and interpretational issues. On the one hand, it is difficult to perform complicated associations between subliminal, rapidly presented stimuli (Atas et al., 2014); on the other hand, this kind of learning would not be of particular interest since it would not add any insightful perspective to the existing frameworks (Kido and Makioka, 2015). Finally, and rather surprisingly, there has been only a single attempt, to the extent of our knowledge, at studying directly the unconscious learning of supraliminal cues governed by simple rules (bottom left corner of Fig. 1A). In this series of experiments from the late 1980s, Lewicki and colleagues found that human participants could learn hidden covariations between the features of different stimuli in the absence of explicit awareness (Lewicki et al., 1987, 1992, 1994). Despite the interest raised by this approach, the interpretation of these studies has been thoroughly criticized (Shanks and Stjohn, 1994) and, even more crucially, their findings have failed the test of independent replication (Hendrickx et al., 1997), leaving the question of the actual existence of this type of unconscious learning unaddressed.

Yet experimental support for the existence of unconscious learning of simple associations between conscious cues would provide decisive responses to enduring criticisms that have been formulated against the existence of unconscious learning. Indeed, Shanks and colleagues have listed four criteria to fulfill in order to provide evidence in favor of unconscious learning: the sensitivity criterion regards the sensitivity of the measures of awareness; the information criterion suggests that the measure of awareness should probe the same information as the experimental task; the immediacy criterion imposes that the tests should be concomitant (or follow immediately) the experimental task; and finally the relevance criterion suggests that the measure of awareness should avoid any irrelevant information. The failure of the current literature in meeting these criteria suggested a substantial lack of evidence in favor of the existence of unconscious learning (Lovibond and Shanks, 2002; Newell and Shanks, 2014; Shanks and Stjohn, 1994).

To address this issue, we developed a simple motion direction discrimination task in which participants were asked to report the motion direction of a colored patch of dots; unbeknownst to them, there was an association between motion direction and 2 out of the 3 possible colors. In 10% of trials, participants were asked to report also the color of the patch together with the motion. In Experiment 1, we tested whether participants were able to learn this association. In Experiment 2, we specifically investigated the extent to which participants were aware of the relevant contingencies, addressing the four criteria suggested by Shanks and colleagues (Newell and Shanks, 2014; Shanks and Stjohn, 1994). In Experiment 3, we studied the role of feedback in such learning. Finally, in Experiment 4, we tested whether learning involved either sensory–motor or sensory–sensory associations.

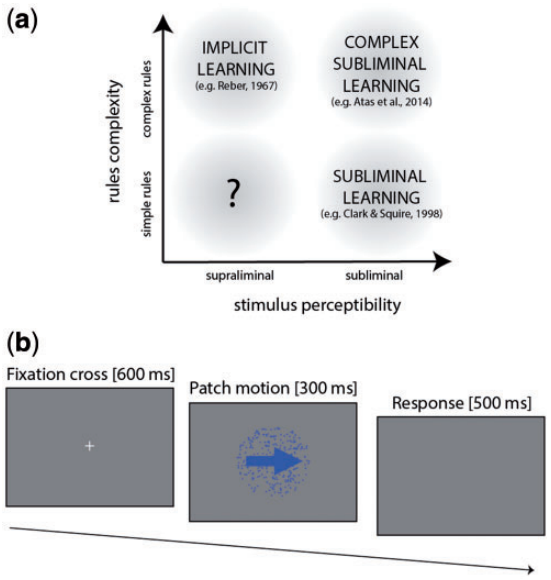

Figure 1. Schematic depiction of the unconscious learning framework (a) and experimental design (b). In the scheme (a), the x-axis represents the stimulus perceptibility (subliminal or supraliminal) and the y-axis represents the rules complexity (simple–complex rules). Within this space, we define four possible categories of unconscious learning paradigms. In the lower part (b), the experimental design is shown: following a fixation cross displayed for 600 ms, a patch of moving dots was displayed for 300 ms. The participants had 500 ms to provide a response.

Experiment 1

Participants

Fourteen healthy participants (7 females, mean age = 24.2 years, SD = 5.63) took part in the first experiment, receiving monetary compensation for their participation. We chose an *a priori* sample size of 14 subjects, since we had no prior data on which to base our initial estimate. In this and all subsequent experiments, we stopped acquiring subjects when we reached the sample size planned before starting the experiments. All of them reported normal or corrected-to-normal vision. All experiments were carried out according to the Declaration of Helsinki and were approved by the Ethics Committee of the Université catholique de Louvain. Written informed consents were obtained from all the participants.

Procedure

The experiment took place in a quiet room, with the participants sitting comfortably on a chair in front of a 19" CRT screen with a 100 Hz refresh rate. The distance between the screen and the chin support was 58 cm. The task was implemented using version 3.0.9 of the Psychotoolbox (Brainard, 1997) in Matlab 7.5 (The MathWorks, Inc. Natick, MA, USA).

At the beginning of each trial, a white cross was displayed at the center of the screen for 600 ms on a gray background (gray levels 0.7). A patch of moving dots was then displayed for 300 ms in the center of the screen, followed by a 500 ms blank screen, after which a new trial began (Fig. 1B). The stimulus was a patch of 2400 dots with a radius of 12°. Each dot was updated every 10 ms. All the dots had a lifetime of three frames, during which they followed a straight line and were then displaced to a new random location. In order to minimize the retinal persistence effect and avoid the perception of the dot trajectory as lines on the screen, each dot was displayed every other frame,

such that there were two interleaved streams of dots displayed alternately. The coherence of the dot motion, i.e. the percentage of dots moving coherently in the same direction, was kept constant during each block.

Participants were instructed to discriminate the motion direction of the dots. They could respond anytime from the patch onset until the end of the trial, by clicking on the left or right mouse buttons with their right hand to indicate a leftward or rightward motion, respectively. Auditory feedback was provided at the end of each trial, signaling a correct (high pitch) or incorrect (low pitch) response. Failure to provide a response was considered as an error.

The experiment consisted of 20 blocks of 60 trials each, and lasted around 45 min altogether (Table 1). The task started with seven blocks of "perceptual training" in which the stimulus consisted exclusively of white dots (gray levels 0.1) in order to determine the level of coherence that will be used for each individual for the rest of the experiment. Indeed, during this training phase, the coherence of the patch was tuned block by block so as to maintain the response accuracy between 70% and 80%. The first block was always performed with 100% coherence patches, but in the following blocks, as soon as accuracy reached 90%, the coherence level was decreased by 15% (or by 20% if accuracy was 100%) in the subsequent block. Conversely, if accuracy fell below 70% in a block, coherence was increased by 10% in the subsequent one. In the following 13 blocks of the "unconscious learning" phase (Table 1), colored patches were shown (red, green, or blue), and motion coherence was kept constant and equal to the coherence value of the last block of the "perceptual training" phase. Crucially, and unbeknownst to participants, in the "unconscious learning" phase, colors and motion direction were associated: one color was always presented in association with leftward motion, another color was always associated with rightward motion, and the third color was equally likely to be associated with rightward and leftward patch motion (Table 2). The color–motion associations were pseudo-randomized across subjects; colors were pseudo-randomly interleaved every three trials (e.g. red–blue–green, green–red–blue, etc.), such that no color appeared more often than twice in a row. Importantly, the association between color and motion was discontinued during the 16th and 17th blocks and restored in the last three blocks (Table 1). Furthermore, to ensure that color information was actually processed by participants, they were asked to report the color of the patch in 10% of the trials, selected randomly.

Table 1. Experimental designs of the four experiments

Color	Motion	Response
Color 1	Right	Right
Color 2	Left	Left
Color 3	R/L	R/L

Participants responded by clicking with the mouse on one of three colored circles displayed on the screen. None of the participants reported explicit awareness of the association between colors and motion direction when questioned about it during debriefing at the end of the experiment.

Data analysis

The behavioral data were analyzed with Generalized Linear Mixed Models (GLMM) implemented in the SAS 9.3 Software (SAS Institute, Cary, NC, USA). We tested two GLMMs, considering accuracy (binary) and reaction times (normal) as dependent variables. In both models, the following independent factors were considered: BLOCK (from 8 to 20), ASSOCIATION (set to 1 if the color was associated with a motion direction, 0 otherwise), COLORS (three levels; included to account for possible effects of the color of the patch on motion discrimination irrespective of its association to a motion direction), and SUBJECT (to account for intersubject differences). Both models were fitted using all the trials, avoiding any preprocessing.

Results

We found that participants' accuracy was higher in the trials in which the color of the dots provided information on the motion direction, as revealed by the factor ASSOCIATION in the GLMM analysis ($F(1,10698) = 13.13$, $P < 0.001$) (Fig. 2). The lack of significant ASSOCIATION × BLOCK interaction prevents us from concluding anything about the dynamics of the learning, but suggests, in accordance with the previous studies (Turk-Browne *et al.*, 2010), that such learning already occurs in the early phase of the experiment (Supplementary Fig. S1). The behavioral advantage induced by the predictive colors disappeared in blocks 16 and 17, during which the contingencies were disrupted, and quickly recovered in the subsequent blocks, after the association had been restored, as confirmed by a GLMM performed on data gathered from blocks 15th to 18th. This analysis revealed a significant effect of the factor ASSOCIATION ($F(1,4098) = 11.68$, $P < 0.001$), and of the BLOCK × ASSOCIATION interaction ($F(4,4098) = 2.75$, $P = 0.026$). A Tukey-corrected *post-hoc* analysis of this interaction revealed better accuracy for predictive colors in block 15th ($t = -3.39$, $P < 0.001$) and 18th ($t = -2.59$, $P = 0.009$), but not in blocks 16th and 17th ($t = -0.48$, $P > 0.250$ and $t = -0.51$, $P > 0.250$). Overall, these analyses confirmed that participants learned the associations between color and motion direction. Regarding the reaction times (RT), only the factor ASSOCIATION was close to be significant ($F(1,10533) = 3.20$, $P = 0.0736$), whereas all other factors and interactions were far from being significant (all $F < 1.8$, $P > 0.10$) (Supplementary Fig. S2). This was expected, given that the motion signal was displayed for only 300 ms and that the participants were provided a very short period of time to respond, in order to emphasize the effect of the color–motion association on accuracy rather than on RT.

Table 2. Associations between color, motion direction, and response in Experiments 1, 2, and 3

	Perceptual training (white dots)	Unconscious learning (colored dots)	Change in the associative structure (colored dots)	Unconscious learning (colored dots)	Explicit tasks (familiarity, generative, questionnaire)	Explicit Blocks (colored dots)
Exp 1	7 blocks	8 blocks	2 blocks	3 blocks		
Exp 2	7 blocks	8 blocks			2 blocks	2 blocks
Exp 3	7 blocks	8 blocks	2 blocks	3 blocks		
Exp 4	8 blocks	8 blocks	2 blocks	2 blocks		

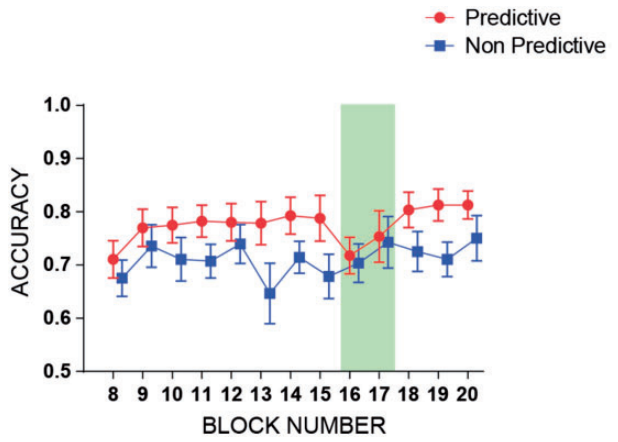

Figure 2. Performance data from Experiment 1. Error bars represent standard errors. The colors represent the two conditions (red: informative color condition, blue: nonpredictive color condition). The violation of the association in block 16th and 17th is marked by the light green rectangle. The performance during the "perceptual training" phase with white patches (blocks 1 to 7) is not displayed.

Discussion

Overall, this first experiment provides evidence in favor of robust learning of the color–motion association. In Experiment 2, we aimed at testing directly participants' awareness of the color–motion direction association by means of two tasks and one questionnaire.

Experiment 2

Participants and procedure

Twenty-three healthy participants (12 females, mean age = 22.73 years, SD = 2.59) participated in the second experiment for monetary compensation. Because we expected negative results in the awareness tasks, we increased the sample size to 24 in order to improve our statistical power. Since one of the subjects did not come on the day of the experiment, we finally acquired data on 23 subjects. All of them reported normal or corrected-to-normal vision. The first part of the experiment (i.e. "perceptual training" and "unconscious learning" phases) was exactly the same as in Experiment 1 (Table 1). The second part was executed immediately after the first one and consisted of three tasks, performed in a pseudo-randomly order by the different subjects. The first task was a generation task consisting of two interleaved types of trials. In the first type of trials (generative color trials), a patch of white dots moved leftward or rightward for 300 ms, while in the second type of trials (generative motion trials), a static but colored patch was displayed for the same amount of time. Participants were asked to associate either a color to the white moving patch, or a motion direction to the colored but static dots; the response was provided by clicking with the mouse on the selected color or motion direction. The entire block was composed of 60 trials (30 for each type). The other task was a familiarity test in which participants were asked to rate, from 1 to 10, the familiarity of displayed patches. All six possible combinations of colors and motion directions were included in this task, which thus also included color–motion combinations to which participants had never been exposed during the experiment. Overall, in this task, 42 trials were performed. The generative and familiarity tasks took about 4 min each. The third test was a short questionnaire with three questions: the first question inquired about any perceived difference between the rightward and leftward motion directions, the second one concerned colors, and the last question asked participants to indicate explicitly whether they had noticed any association between color and motion direction. Finally, participants were told about the association between colors and motion directions, and were asked to perform two additional blocks of the discrimination task while being now explicitly aware of the association.

The same analyses as in Experiment 1 were performed twice on two distinct datasets: once on all the subjects ($n = 23$), and then only on those who did not provide a correct response in the questionnaire about the color/motion association ("implicit" group, $n = 18$). Given that the results of these analyses unveiled a lack of effect, we computed the Bayes Factor (BF) to estimate the likelihood of the null hypothesis being true (Smith, 2001). The BF can be used as an alternative way to test statistical hypothesis. It relies on the estimation of the probability of a statistical model (or hypothesis) given the observed data. One major advantage of this approach is that it allows researchers to estimate the validity of the null hypothesis, in comparison to alternative hypotheses. Indicatively, a BF between 0.3 and 3 suggests a lack of sensitivity. A BF below 0.3 or above 3 provides strong evidence in support for the alternative hypothesis, or for the null hypothesis, respectively. In order to compute the BF, we compared the Bayesian Information Criterion (BIC) estimated from each model with and without the explanatory variable (Masson, 2011; Smith, 2001).

Results

We confirmed that the discrimination accuracy was higher in trials with predictive colors than with the control color (Fig. 3) (ASSOCIATION: $F(1,10664) = 4.58$, $P < 0.001$). We also found a progressive increase in accuracy across blocks (BLOCK: $F(7,154) = 4.58$, $P < 0.001$) but no significant ASSOCIATION × BLOCK interaction ($F(7,10664) = 0.69$, $P > 0.250$) (see also Supplementary Fig. S2). Similar results were obtained when restricting the analysis to the "implicit group" only ($n = 18$; ASSOCIATION: $F(1,8344) = 17.44$, $P < 0.001$; BLOCK: $F(7,119) = 3.23$, $P = 0.003$, interaction $F(7,8344) = 0.84$, $P > 0.250$). Regarding the RT (Supplementary Fig. S1), only the factor ASSOCIATION was significant, with faster responses being associated with predictive colors (ASSOCIATION: $F(1,10499) = 25.93$, $P < 0.001$; BLOCK: $F(7,154) = 1.23$, $P > 0.250$; interaction $F(7,10499) = 1.00$, $P > 0.250$; implicit group, $n = 18$: ASSOCIATION: $F(1,8200) = 10.77$, $P = 0.001$; BLOCK: $F(7,119) = 1.21$, $P > 0.250$; interaction $F(7,8200) = 0.89$, $P > 0.250$).

We compared the accuracy and RT of all the participants in the last two blocks of the "unconscious learning" phase (blocks 14th and 15th) with the data obtained in the "explicit blocks" performed at the end of the experiment (E1 and E2, Fig. 3), after the color–motion association had been explicitly revealed to the subjects (Table 1). As expected, we found a significant effect of the factor ASSOCIATION on accuracy ($F(1,5426) = 145.95$, $P < 0.001$), but not of the factor accounting for the awareness of the subjects (EXPLICITNESS: $F(1,22) = 1.59$, $P = 0.220$); however, the interaction between these two factors was highly significant ($F(1,5426) = 45.86$, $P < 0.001$). Similarly, the analysis on RT revealed an effect of all the factors and their interaction (ASSOCIATION: $F(1,5326) = 168.94$, $P < 0.001$; EXPLICITNESS: $F(1,22) = 17.41$, $P < 0.001$; interaction: $F(1,5326) = 89.31$, $P < 0.001$). These changes in both RT and accuracy following

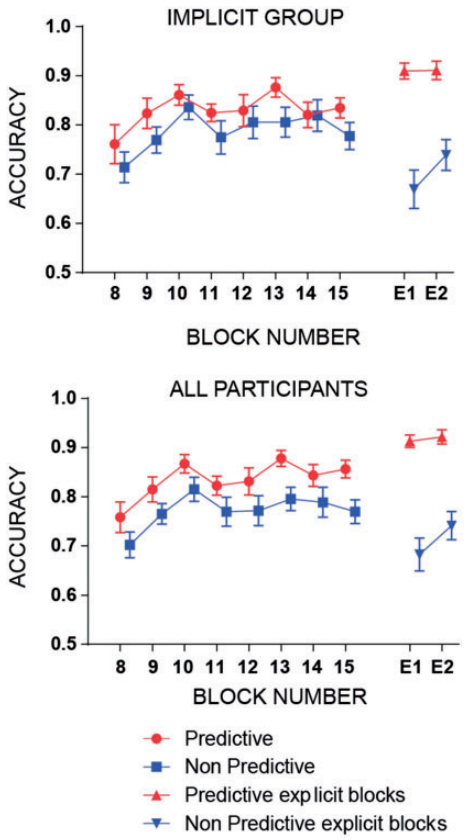

Figure 3. Accuracy data from Experiment 2. In the upper part, the data from the group of subjects who did not provide the correct associations in the questionnaire is shown (implicit group, $n = 18$), whereas all the participants are included in the data shown in the lower panel ($n = 23$). Error bars represent standard errors. E1 and E2 are two blocks in which participants were instructed about the associations. The data from the "perceptual training" phase with white patches (from 1 to 7) is not displayed in the figure.

awareness of the association suggests a drastic change in strategy, in which participants started presumably to base their decision primarily on the color information.

Concerning the questionnaire, only 5 subjects out of 23 explicitly reported the color–motion association. Another subject reported only one correct color–motion association, and two other subjects reported incorrect color–motion associations. All other subjects reported having noticed no systematic association whatsoever.

Regarding the generative tasks, in trials in which participants were asked to associate a motion direction to a color (Fig. 4A), we considered as dependent variables the motion direction chosen by the participants, and, as factors, the COLORS and TASK-ORDER (categorical variable, accounting for whether the task was performed before or after the questionnaire). We found no significant main effects for either group (all subjects, $n = 23$: COLORS: $F(2,659) = 0.29$, $P > 0.250$; TASK-ORDER: $F(1,659) = 0.29$, $P > 0.250$; "implicit group," $n = 18$: COLORS: $F(2,48) = 0.12$, $P > 0.250$; TASK-ORDER: $F(1,465) = 0.79$, $P = 0.3752$), but found a significant interaction for the group including all the subjects (all subjects: $F(2,659) = 3.07$, $P = 0.0474$), suggesting that, after the questionnaire, subjects were more likely to associate the informative colors to the correct motion direction, as revealed by a significant *post-hoc* analysis (difference between left-motion associated color and right-motion associated color: $t = 4.03$, Tukey–Kramer corrected $P < 0.001$; between second informative color and noninformative color $t = 6.06$, Tukey–Kramer corrected $P < 0.001$). The significant interaction was, however, not observed in the "implicit group" ($F(2,465) = 0.53$, $P > 0.250$). To confirm these negative findings in the implicit group, we computed a BF by comparing the BIC obtained from each model with and without the explanatory variable COLORS (Masson, 2011; Smith, 2001). The results confirm a lack of effect of the COLORS factor for the implicit group (implicit group: BF = 33.11, $P < 0.03$). The second generative task (Fig. 4B), in which participants associated a color to a given motion direction, revealed no effects in any factors or interaction for both groups (all the subjects: MOTION-DIRECTION: $F(1,42) = 0.58$, $P > 0.250$; TASK-ORDER: $F(1,643) = 0.23$, $P > 0.250$); interaction: $F(1,643) = 0.91$, $P > 0.250$); implicit group: MOTION-DIRECTION: $F(1,32) = 1.32$, $P = 0.2599$; TASK-ORDER: $F(1,503) = 0.01$, $P > 0.250$); interaction: $F(1,503) = 2.10$, $P = 0.1480$). The BF confirmed the lack of result for the factor MOTION-DIRECTION and its interaction for both groups (all subjects: BF = 20.08, $P < 0.05$, implicit group: BF = 90.01, $P < 0.02$).

Considering the familiarity task (Fig. 4C), we did not find any significant effect for the factors ASSOCIATION (correct association, incorrect association or control color; $F(2,48) = 0.14$, $P > 0.250$), TASK-ORDER ($F(1,688) = 2.09$, $P = 0.1491$), and their interaction ($F(2,688) = 0.46$, $P > 0.250$) for the implicit group ($n = 18$). The BF confirmed the lack of effect for the ASSOCIATION factor in the model (BF = 38.11, $P < 0.03$). When considering all subjects, we found a significant effect of the interaction between the factors ($F(2,881) = 3.31$, $P = 0.0370$), but nor a main effect of the task order, neither of the association factor (ASSOCIATION: $F(2,63) = 1.14$, $P > 0.250$; TASK-ORDER: $F(1,881) = 0.84$, $P > 0.250$). This interaction shows a significant difference between the first informative color and the noninformative color, as revealed by a *post-hoc* analysis ($t = 3.86$, Tukey–Kramer corrected $P < 0.002$).

Moreover, to confirm that all the tests measured the same variable (i.e. awareness), we correlated the results of the generative and familiarity tasks. In order to do so, we computed the Euclidean distance from the rates (or proportions of answers in the generative tasks) provided by each subject to the optimal "explicit" behavior, such that small distances would reflect more awareness of the associations. The results confirmed a correlation between the two generative tasks (Pearson correlation, $r = 0.53$, $P < 0.01$) and between the familiarity task and the second generative task (assigning color to motion, $r = 0.5823$, $P < 0.01$), but not between familiarity and the first generative task (assigning motion to color, $r = 0.1648$, $P > 0.25$). Altogether, these correlational analyses confirmed that all the tests measured the awareness of the learnt associations (Supplementary Fig. S3).

Discussion

In Experiment 2, we found that when restricting the analyses to the group of subjects that failed to report the association between color and motion, the learning of this association was still robust in spite of the fact that none of the awareness tasks showed significant results. This demonstrates that even though some participants gained explicit awareness of the association, learning took place in the absence of awareness for most of them.

The difference between the last two blocks of the implicit phase and the two explicit blocks highlighted a change in the

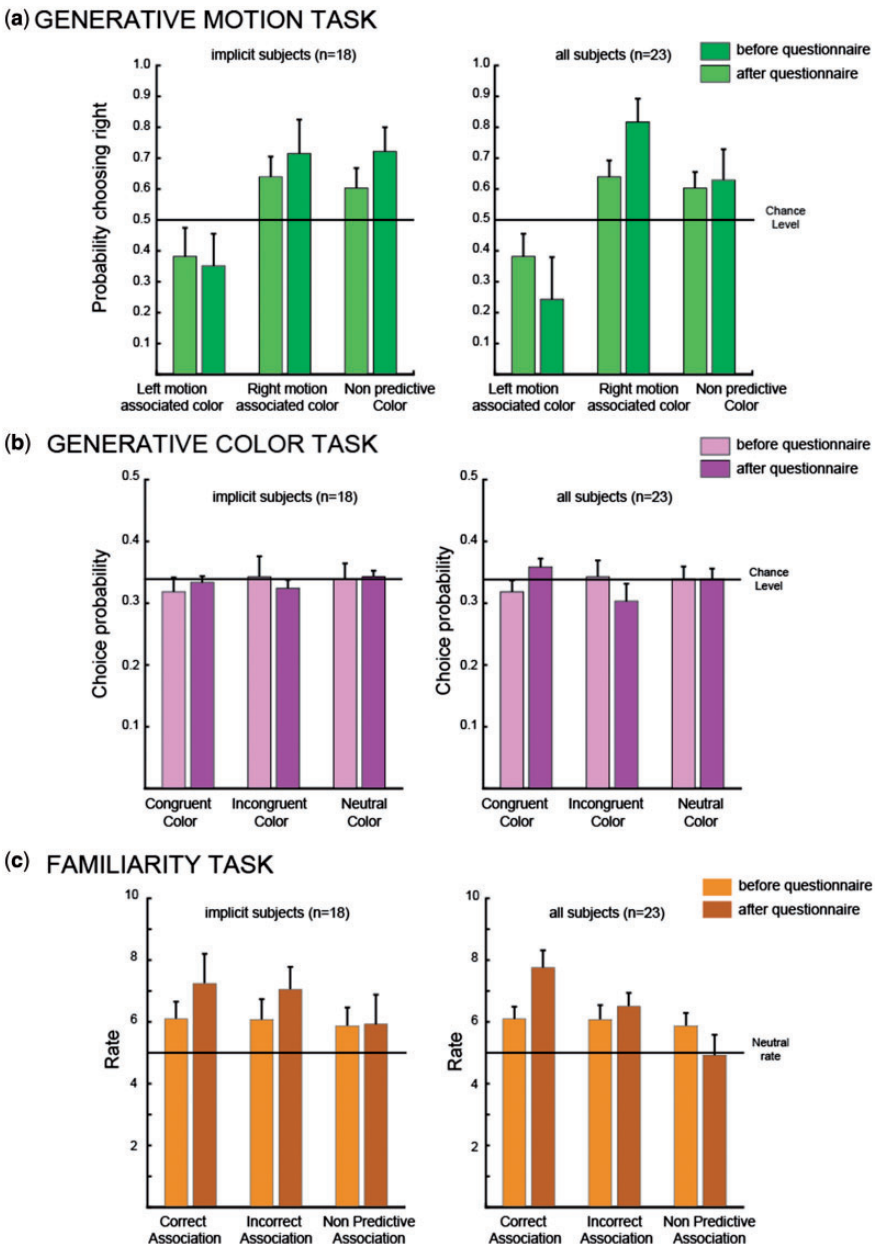

Figure 4. Results from the awareness testing tasks. The data from the participants who did not respond correctly to the questionnaire (implicit group, $n = 18$) is shown on the left, whereas all participants are included in the dataset shown on the right ($n = 23$). Participants were split based on whether they performed the awareness testing tasks before (light colors) or after (darker colors) the questionnaire. Results from the generative motion task (a): on the y-axis, the probability of choosing rightward motion is shown. The results from the generative color task are displayed in the middle part (b): values on the y-axis represent the probability of choosing either of the three color conditions displayed along the x-axis (congruent, incongruent, or neutral color). In the last part (c) the results from the familiarity task are shown: on the y-axis, the average familiarity ratings associated with the three different types of patches are shown (correct predictive association, incorrect association, or nonpredictive association).

strategy adopted by the participants: when the association became fully explicit, participants seem to focus mostly on the color feature rather than on the motion. Such difference strongly suggests that if the participants had explicit knowledge of the association during the main experiment, their behavior should have been similar to that exhibited in the explicit blocks. These results thus provide strong evidence for the unconscious nature of the learning.

In Experiment 3, we attempted to gain insight into the basic learning mechanisms at play in our task. We wondered whether participants learned the association through reinforcement or Hebbian learning. To address this question, we investigated the role of feedback in the learning of the color-motion association.

Experiment 3

Participants and procedure

Fourteen healthy participants (9 females, mean age = 23.42 years, SD = 1.74) participated in the third experiment for

monetary compensation. All of them reported normal or corrected-to-normal vision. The experimental design was the same as in Experiment 1, except that no auditory feedback was provided during the "unconscious learning" phase of the experiment (Table 1). Therefore, we used the same sample size as in the first experiment. Auditory feedback was still provided during the first 7 blocks of training with the white dots patches in order to obtain a level of response accuracy between 70% and 80% as in Experiment 1.

Results

As in the previous experiment, we analyzed accuracy and RT in two GLMMs. Regarding accuracy (Fig. 5), we did not find any significant effect: ASSOCIATION ($F(1,10697) = 2.41$, $P = 0.1204$), BLOCK: $F(12,156) = 0.87$, $P > 0.250$, COLORS: $F(2,25) = 1.29$, $P > 0.250$, BLOCK × ASSOCIATION: $F(12,10697) = 1.15$, $P > 0.250$). In order to test specifically the lack of effect of the factor ASSOCIATION, we computed the BF on the basis of the BIC obtained from the models with and without this factor. This analysis confirmed the lack of effect of the factor ASSOCIATION on the response accuracy (BF $= 2.8478 \times 10^{11}$, $P < 0.0001$). The very large value of the BF, along with the highly significant estimated P-value, confirmed that the sample size provided enough statistical power to properly test the hypothesis. Concerning the RT (Supplementary Fig. S1), we found a significant effect of COLORS ($F(2,25) = 4.04$, $P = 0.0301$), while the other effects were not significant (ASSOCIATION $F(1,10254) = 0.05$, $P > 0.250$; BLOCK: $F(12,156) = 0.85$, $P > 0.250$; BLOCK × ASSOCIATION $F(12,10254) = 1.35$, $P = 0.1850$)). The color effect revealed that overall the subjects were slower in detecting the blue color, irrespective of its association with the motion direction.

We then compared the response accuracy in Experiments 1 and 3, by means of a GLMM with COLORS, ASSOCIATION, BLOCK, SUBJECT, and EXPERIMENT as factors. Interestingly, the results revealed a significant effect of the factor ASSOCIATION ($F(1,21615) = 33.68$, $P < 0.001$), BLOCK ($F(12,156) = 1.91$, $P = 0.0375$), EXPERIMENT ($F(1,21615) = 42.44$, $P < 0.001$) and the interaction EXPERIMENT × ASSOCIATION ($F(1,21615) = 17.55$, $P < 0.001$), thus confirming the importance of the auditory feedback not only in performing the task, but also in the unconscious learning of the association.

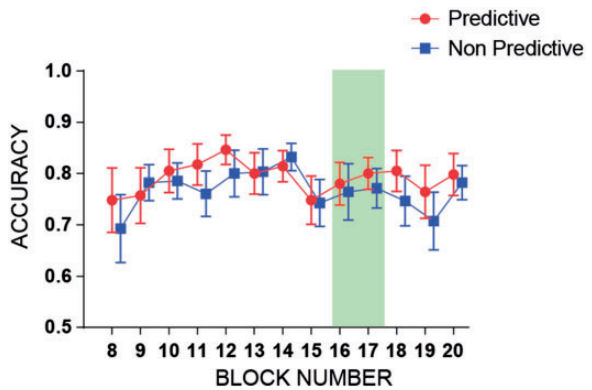

Figure 5. Performance data from Experiment 3. The informative/non-informative colors are represented in red/blue. Error bars represent standard errors. The violation of the association in block 16th and 17th is marked by the light green rectangle. The data from the "perceptual training" phase with white patches (from 1 to 7) is not displayed.

Discussion

Experiment 3 indicates that when no feedback on response accuracy is provided, learning fails to take place. This suggests that this type of unconscious learning relies on reinforcement learning mechanisms (Dayan and Balleine, 2002; Sutton and Barto, 1998), rather than on Hebbian-like associative learning between concurrent features of the stimuli (Munakata and Pfaffly, 2004). Since reinforcement learning is usually studied in the context of stimulus-response associations (Dayan and Balleine, 2002), these findings suggested that participants may have learned to associate the color of the stimulus with the response dictated by the corresponding motion direction, as opposed to associating directly the color with the motion. We tested this hypothesis in Experiment 4.

Experiment 4

Participants and procedure

Twenty-four healthy participants (14 females, mean age = 27.04, SD = 4.78) joined the fourth experiment for monetary compensation. All of them reported normal or corrected-to-normal vision. The sample size was chosen from an expected effect size of 0.4676% (estimated from Experiment 1) and aiming for a power of 0.80 and an alpha of 0.05. The experimental design was similar to that of Experiment 1, except as detailed below.

The "perceptual training" phase with the white dots lasted for 8 blocks, followed by 12 blocks of "unconscious learning" phase with colored patches and fixed coherence levels (Table 1). Auditory feedback was provided trial by trial in every block. The motion of the dots was directed upward or downward, and the participants were instructed to respond left or right to indicate the motion direction. The rule linking the response to the motion direction changed in every block, so that if the upward motion was associated with the left response in one block, the left response was associated with the downward motion in the subsequent block. This alternation of the rules was reminded to the subjects both verbally by the experimenter and visually at the beginning of each block by displaying a message for 5 s on the computer screen. Importantly, the subjects were randomly separated in two groups: in the first group, the association between color and motion was kept constant during the whole experiment, while in the second group, it was the association between color and response which was kept constant (Table 3). In this way, only one association (either color–motion direction or color–response side) was maintained during the whole experiment. Similar to Experiment 1, these associations were discontinued in blocks 17 and 18 (Table 1).

We performed two GLMMs considering as dependent variables either the accuracy or the RT. The only main difference with the previous analyses was the independent factor named GROUP, which indicated to which group the subject was assigned (either color–motion association or color–response association).

Results

The first GLMM performed on accuracy (Fig. 6) revealed a significant effect of the factors BLOCK ($F(11,198) = 4.07$, $P < 0.001$), ASSOCIATION ($F(1,14098) = 8.16$, $P = 0.0043$), and their interaction ($F(11,14098) = 3.12$, $P < 0.001$). A *post-hoc* analysis of the interaction revealed a significant effect of ASSOCIATION for blocks 9, 12, 13, 15, and 16 (all $t < -2.10$, Tukey–Kramer

Table 3. Associations between color, motion direction, and response in Experiment 4

Color–motion group		
Block i		
Color	Motion	Answer
Color 1	**Upward**	Right
Color 2	**Downward**	Left
Color 3	U/D	R/L
Block i + 1		
Color	Motion	Answer
Color 1	**Upward**	Left
Color 2	**Downward**	Right
Color 3	U/D	R/L
Color–response group		
Block i		
Color	Motion	Answer
Color 1	Upward	**Right**
Color 2	Downward	**Left**
Color 3	U/D	R/L
Block i + 1		
Color	Motion	Answer
Color 1	Downward	**Right**
Color 2	Upward	**Left**
Color 3	U/D	R/L

corrected $P < 0.03$). Surprisingly, the *post-hoc* analysis revealed also a significant difference in block 18, the second block in which the rules were violated, but highlighting a higher accuracy for the noninformative color ($t = 2.11$, $P = 0.0347$). In the last two blocks, when the rules were restored, no difference emerged from the *post-hoc* analysis (block 19: $t = -0.88$, $P = 0.3787$, block 20: $t = 0.26$, $P > 0.250$). Crucially, no difference between the two groups emerged (GROUP: $F(1,14098) = 0.00$, $P > 0.250$; GROUP × ASSOCIATION $F(1,14098) = 1.98$, $P = 0.1592$; GROUP ×ASSOCIATION × BLOCK $F(11,14098) = 0.66$, $P = 0.7751$), indicating that both groups learnt equally well. The BF comparing the model with and without the factor GROUP confirmed this lack of significant effect ($BF = 2.1138 \times 10^{18}$, $P < 0.0001$). Regarding the analysis of the RT (Supplementary Fig. S1), only the BLOCK factor and the BLOCK × GROUP interaction revealed a significant effect (BLOCK: $F(11,198) = 10.45$, $P < 0.001$, BLOCK × GROUP $F(11,13832) = 1.83$, $P = 0.0438$); no effect of the factor ASSOCIATION ($F(1,13832) = 1.83$, $P = 0.1758$) or GROUP ($F(1,13832) = 2.71$, $P = 0.0996$) was revealed. Regarding the BLOCK × GROUP interaction, none of the pairwise *post-hoc* comparisons showed a significant difference between the groups, suggesting that the effect was driven by different trends in the two groups, with the first group reducing its reaction time abruptly in the 5th block (Supplementary Fig. S1). Finally, we compared the results from Experiments 1 and 4, adding to the GROUP variable a level accounting for the data from Experiment 1. GROUP was thus composed of three levels: two for each group of Experiment 4 and one for Experiment 1. No difference between the groups was revealed by the GLMM (GROUP × ASSOCIATION: $F(2,28043)$, $P > 0.25$), as confirmed by the BF analysis performed comparing the BIC of the model with and without the interaction GROUP × ASSOCIATION ($BF = 1.406 \times 10^7$, $P < 0.001$).

Discussion

In contrast to our expectation, we found that both the sensory–sensory and the sensory–motor associations were learnt equally well, with no difference between the groups. This suggests that the mechanisms involved in this type of associative learning are general rather than modality-specific, and that the events whose association gets unconsciously learned do not have to be perfectly concurrent in time, since the response followed the color by a few hundreds of milliseconds.

General Discussion

In the present study, we investigated whether participants can learn unconsciously a direct association between supraliminal features of task-relevant stimuli. The results showed robust and reliable learning of the association. Importantly, in Experiment 2 we directly tested the degree of awareness of such learning in light of the four criteria suggested by Shanks and Newell in 1994 (Dawson and Reardon, 1973; Newell and Shanks, 2014; Shanks and Stjohn, 1994). Briefly, these criteria require that the assessment of awareness should be devoid of biasing factors not relevant for the behavioral measure (reliability), that they should be performed immediately after the experiment (immediacy), and under optimal retrieval conditions (sensitivity; Newell and Shanks, 2014). The fourth crucial criterion is the relevance criterion (Shanks and Stjohn, 1994), which stipulates that the test of awareness should target the very same information that drives changes in behavior. In our design, given the simple nature of the learned association, it is easy to fulfill both the reliability and relevance criteria. Whereas in standard implicit learning paradigms, participants may potentially achieve success by exploiting information that the test of awareness fails to probe, in our case, the simplicity of the contingency excludes this possibility. Indeed, only the learning of the association between the color of the dots and the direction of their motion can lead to the improvement in accuracy that we observed. Further, we extensively probed participants' conscious knowledge of this association directly by means of generation and familiarity tasks, and through a questionnaire. Regarding the sensitivity criterion, our Bayesian approach made it possible to convincingly conclude [BF > 30 (Dienes, 2011)] that our null findings can be interpreted as offering support for the absence of differences rather than as resulting from a lack of sensitivity (Vadillo *et al.*, 2015). Finally, regarding the immediacy criterion, we administered the tests as soon as the learning phase ended, so reducing the effect of interference or forgetting as much as possible. It is noteworthy that, during the generative task, the group who did not report the correct association in the questionnaire performed at random when coupling the colors with the motion directions. This is a substantive finding because it suggests that the color–motion association is used only in the narrow context of the task in which it was learned and cannot be transferred to different task-sets (Graf and Schacter, 1985; Shanks *et al.*, 1997).

We believe that our study provides the first demonstration of unconscious learning of simple associations, thereby filling the gap in the theoretical framework illustrated in Fig. 1. Previous studies in the framework of delay conditioning (Clark and Squire, 1998; Clark *et al.*, 2002), fear conditioning (Maren, 2001) or in the context of the relationship between unconscious processing and perceptual load (Bahrami *et al.*, 2008; Carmel *et al.*, 2007) could be viewed as providing already indirect evidence for unconscious learning of supraliminal stimuli. However, these studies did not test systematically the

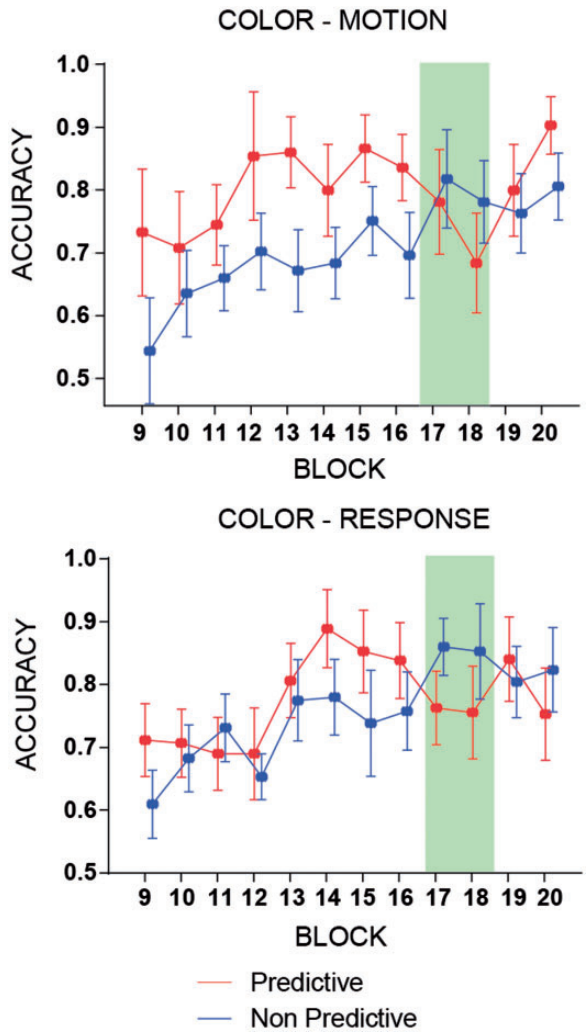

Figure 6. Performance in Experiment 4. In the upper panel, the data from the group in which the *color–motion* association was kept constant ($n = 12$) is shown, and in the lower panel the data from the group in which the *color–response* association was kept constant is represented ($n = 12$). The red and blue dots correspond to the informative and noninformative colors, respectively. Error bars represent standard errors, and violation of the association in block 16th and 17th is marked by the light green rectangle. The data from the "perceptual training" phase with white patches (from 1 to 8) is not displayed in the figure.

awareness of the learned association, thus failing to address the criteria discussed above (Shanks and Stjohn, 1994). One of the possible reasons for this lack of previous demonstration of the unconscious learning of simple associations could be that designing an experiment in which the association between perceivable stimulus features is unconsciously learnt was quite challenging. On the one hand, the predictive cue should be perceived and actively processed by the participants in order to affect their behavior (Jiang and Chun, 2001; Jimenez and Mendez, 1999); on the other hand, the contingency with the other feature should remain implicit. The balance between these two extremes was difficult to obtain. One important aspect of our task is that the color had to be processed actively by the participants because of the secondary task, which consisted in reporting the color of the patch in 10% of the trials. Nevertheless, participants failed to perceive explicitly the association, possibly because of what can be defined as "a change of narrative": the secondary task related to the predictive cue is a sufficiently convincing justification for the presence of the colors in the task, such that participants do not have to search for an explanation, which would eventually lead them to figure out the association (Shanks, 2003). Future experiments should further investigate whether this interpretation holds true. Furthermore, the presence of the secondary color task allows us to exclude inattentional amnesia as an alternative interpretation (Wolfe, 1999), since it compelled the participants to direct actively their attention to the colors during the task.

Whereas in Experiments 1 and 2, we provide evidence in favor of the learning without awareness of the color–motion association, in Experiments 3 and 4 we investigate the possible mechanisms involved in this type of learning. Specifically, Experiment 3 showed that auditory feedback was necessary for learning to occur, thus evoking reinforcement learning (Niv, 2009), while Experiment 4 showed that both the color–motion and the color–response associations were learnt. Feedback can, in some cases, be processed as a reward signal and induce a phasic dopaminergic response (Hyman *et al.*, 2006) which would then reinforce the circuits that link the color with the motion features or the color with the response representation. This interpretation concurs with a recent review on the pharmacology of implicit learning suggesting a similar link with dopaminergic systems (Uddén *et al.*, 2010).

Conclusion

In conclusion, we provide a novel and robust experimental design that can be used to investigate unconscious associative learning, and we have begun to decipher its basic mechanistic features. However, many important questions remain unanswered. For instance, we found a large inter individual variability in both the magnitude of the learning and the level of awareness of the association. Exploring the extent to which attention, working memory capacity or cognitive control (Stillman *et al.*, 2014) are involved in this variability is an important goal for further research.

Supplementary data

Supplementary data is available at *Neuroscience of Consciousness Journal* online.

Acknowledgements

This work was performed at the Institute of Neuroscience (IoNS) of the Université catholique de Louvain (Brussels, Belgium); it was supported by grants from the ARC (Actions de Recherche Concertées, Communauté Française de Belgique), from the Fondation Médicale Reine Elisabeth (FMRE), and from the Fonds de la Recherche Scientifique (FNRS–FDP) to E.O. and A.Z. A.A. is a Research Fellow at the FNRS and A.Z. is a Senior Research Associate supported by INNOVIRIS.

Conflict of interest statement: None declared.

References

Atas A, Faivre N, Timmermans B *et al*. Nonconscious learning from crowded sequences. *Psychol Sci* 2014;**25**:113–9. doi:10.1177/0956797613499591

Bahrami B, Carmel D, Walsh V *et al*. Unconscious orientation processing depends on perceptual load. *J Vis* 2008;**8**:12.1–0. doi:10.1167/8.3.12

Berry DC, Broadbent DE. On the relationship between task performance and associated verbalizable knowledge. *Quart J Exp Psychol Section A* 1984;**36**:37–41. doi:10.1080/14640748408402156

Brainard DH. The psychophysics toolbox. *Spat Vis* 1997;**10**:433–6. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9176952. Last access 2016, August 17.

Carmel D, Saker P, Rees G *et al*. Perceptual load modulates conscious flicker perception. *J Vis* 2007;**7**:1–13. doi:10.1167/7.14.14

Clark RE, Manns JR, Squire LR. Classical conditioning, awareness, and brain systems. *Trends Cogn Sci* 2002;**6**:524–31.

Clark RE, Squire LR. Classical conditioning and brain systems: the role of awareness. *Science* 1998;**280**:77–81. doi:10.1126/science.280.5360.77

Dawson ME, Reardon P. Construct validity of recall and recognition postconditioning measures of awareness. *J Exp Psychol* 1973;**98**:308–15. doi:10.1037/h0034372

Dayan P, Balleine BW. Reward, motivation, and reinforcement learning. *Neuron* 2002;**36**:285–98.

Dienes Z. Bayesian versus orthodox statistics: which side are you on? *Perspect Psychol Sci* 2011;**6**:274–90. doi:10.1177/1745691611406920

Graf P, Schacter DL. Implicit and explicit memory for new associations in normal and amnesic subjects. *J Exp Psychol Learn Mem Cogn* 1985;**11**:501–18. doi:10.1037/0278-7393.11.3.501

Hendrickx H, De Houwer J, Baeyens F *et al*. Hidden covariation detection might be very hidden indeed. *J Exp Psychol Learn Mem Cogn* 1997;**23**:201–20. doi:10.1037/0278-7393.23.1.201

Hyman SE, Malenka RC, Nestler EJ. Neural mechanisms of addiction: the role of reward-related learning and memory. *Ann Rev Neurosci* 2006;**29**:565–98. doi:10.1146/annurev.neuro.29.051605.113009

Jiang Y, Chun MM. Selective attention modulates implicit learning. *Q J Exp Psychol A* 2001;**54**:1105–24. doi:10.1080/713756001

Jimenez L, Mendez C. Which attention is needed for implicit sequence learning? *J Exp Psychol Learn Mem Cogn* 1999;**25**:236–59.

Kido K, Makioka S. Serial order learning of subliminal visual stimuli: evidence of multistage learning. *Front Psychol* 2015;**6**:1–13. doi:10.3389/fpsyg.2015.00076

Lewicki P, Czyzewska M, Hoffman H. Unconscious acquisition of complex procedural knowledge. *J Exp Psychol Learn Mem Cogn* 1987;**13**:523–30. doi:10.1037/0278-7393.13.4.523

Lewicki P, Hill T, Czyzewska M. Nonconscious acquisition of information. *Am Psychol* 1992;**47**:796–801. doi:10.1037/0003-066X.47.6.796

Lewicki P, Hill T, Czyzewska M. Nonconscious indirect inferences in encoding. *J Exp Psychol Gen* 1994;**123**:257–63. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7931091. Last access 2016, August 17.

Lovibond PF, Shanks DR. The role of awareness in Pavlovian conditioning: empirical evidence and theoretical implications. *J Exp Psychol Anim Behav Process* 2002;**28**:3–26. doi:10.1037/0097-7403.28.1.3

Maren S. Neurobiology of Pavlovian fear conditioning. *Annu Rev Neurosci* 2001;**24**:897–931. doi:10.1146/annurev.neuro.24.1.897

Masson MEJ. A tutorial on a practical Bayesian alternative to null-hypothesis significance testing. 2011;679–90. doi:10.3758/s13428-010-0049-5

Munakata Y, Pfaffly J. Hebbian learning and development. *Developmental Science* 2004;**7**:141–8.

Newell BR, Shanks DR. Unconscious influences on decision making: a critical review. *Behav Brain Sci* 2014;**37**:1–19. doi:10.1017/S0140525X12003214

Nissen MJ, Bullemer P. Attentional requirements of learning: evidence from performance measures. *Cogn Psychol* 1987;**19**:1–32. doi:10.1016/0010-0285(87)90002-8

Niv Y. Reinforcement learning in the brain. *J Mathe Psychol* 2009;**53**:139–54. doi:10.1016/j.jmp.2008.12.005

Olsson A, Phelps EA. Learned fear of "unseen" faces after Pavlovian, observational, and instructed fear. *Psychol Sci* 2004;**15**:822–8. doi:10.1111/j.0956-7976.2004.00762.x

Reber AS. Implicit learning of artificial grammars. *J Verb Learn Verb Behav* 1967a;**6**:855–63. doi:10.1016/S0022-5371(67)80149-X

Reber AS. Implicit learning of artificial grammars. 1967b; http://philpapers.org/rec/REBILO (17 August 2016, date last accessed).

Saffran JR, Johnson EK, Aslin RN *et al*. Statistical learning of tone sequences by human infants and adults. *Cognition* 1999;**70**:27–52. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10193055. Last access 2016, August 17.

Shanks DR. Attention and awareness in "implicit" sequence learning. In: Jiménez, L. (ed), *Attention and Implicit Learning. Advances in Consciousness Research*, Vol. **48**, Amsterdam: John Benjamins Publishing, 2003, 11–42.

Shanks DR, Johnstone T, Staggs L. Abstraction processes in artificial grammar learning. *Q J Exp Psychol A* 1997;**50**:216–52. doi:10.1080/713755680

Shanks D, Stjohn M. Characteristics of Dissociable Human Learning-Systems. *Behav Brain Sci* 1994;**17**:367–95. doi:10.1017/S0140525X00035032

Smith AFM, Bernardo JM. Bayesian theory. *Meas Sci Technol*, 2001;**12**(2):221.

Stillman CM, Feldman H, Wambach CG *et al*. Dispositional mindfulness is associated with reduced implicit learning. *Conscious Cogn* 2014;**28**: 141–50. doi:10.1016/j.concog.2014.07.002

Sutton RS, Barto AG. Reinforcement learning: an introduction. *IEEE Trans Neural Netw* 1998;**9**:1054. doi:10.1109/TNN.1998.712192

Turk-Browne NB, Junge JA, Scholl BJ. The automaticity of visual statistical learning. *J Exp Psychol: General* 2005;**134**:552–64. doi:10.1037/0096-3445.134.4.552

Turk-Browne NB, Scholl BJ, Chun MM *et al*. Neural evidence of statistical learning: efficient detection of visual regularities without awareness. *J Neurosci* 2010;**21**:1934–45. doi:10.1162/jocn.2009.21131.Neural

Uddén J, Folia V, Petersson KM. The neuropharmacology of implicit learning. *Curr Neuropharmacol* 2010;**8**:367–81. doi:10.2174/157015910793358178

Vadillo MA, Konstantinidis E, Shanks DR. Underpowered samples, false negatives, and unconscious learning. *Psychon Bull Rev*. 2016;**23**:87–102.

Wolfe JM. Inattentional amnesia Wolfe Jeremy M. In: Coltheart V (ed), *Fleeting Memories: Cognition of Brief Visual Stimuli*, Cambridge, MA: MIT Press, 1999.